Quantum Nonlocality in Weak-Thermal-Light Interferometry

Mankei Tsang*

Department of Electrical and Computer Engineering, National University of Singapore, 4 Engineering Drive 3, Singapore 117583
Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117551
Center for Quantum Information and Control, University of New Mexico, MSC07–4220, Albuquerque, New Mexico 87131-0001, USA

In astronomy, interferometry of light collected by separate telescopes is often performed by physically bringing the optical paths together in the form of Young's double-slit experiment. Optical loss severely limits the efficiency of this so-called direct detection method, motivating the fundamental question of whether one can achieve a comparable performance using separate optical measurements at the two telescopes before combining the measurement results. Using quantum mechanics and estimation theory, here I show that any such spatially local measurement scheme, such as heterodyne detection, is fundamentally inferior to coherently nonlocal measurements, such as direct detection, for estimating the mutual coherence of bipartite thermal light when the average photon flux is low. This surprising result reveals an overlooked signature of quantum nonlocality in a classic optics experiment.

The basic goal of stellar interferometry is to retrieve astronomical information from the mutual coherence between optical modes collected by telescopes [1–3]. The imaging resolution increases with the distance between the collected optical modes called the baseline, motivating the development of long-baseline stellar interferometry using light collected from a telescope array [2,3]. The standard method of stellar interferometry in the optical regime is called direct detection, which coherently combines the optical paths in the form of the classic Young's double-slit experiment, but its efficiency suffers from decoherence in the form of accumulating optical loss along the paths as the baseline is increased. To avoid optical loss, an alternative method is to perform separate heterodyne detection at the two telescopes, before combining the measurement results via classical communication and data processing [2,3]. In quantum information theory, direct detection can be classified as a nonlocal measurement scheme, which requires joint quantum operations on the two optical modes, while heterodyne detection is a local measurement scheme, which does not require quantum coherence between the separate detectors [4,5]. Townes has previously analyzed the quantum noises in direct and heterodyne detection and concluded that direct detection is superior at high optical frequencies and heterodyne detection is superior at low frequencies [3,6]. Heterodyne detection is, however, only one example of local measurements, and it remains a fundamental and important question whether any other local measurement can perform as well as nonlocal measurements while not suffering from decoherence.

The main purpose of this Letter is to prove that, in the case of weak thermal light, *any* local measurement scheme must be significantly inferior to a nonlocal one for the estimation of the mutual coherence according to quantum mechanics. This is a surprising result in quantum metrology, since the disadvantage of local measurements does not otherwise occur for coherent states at any strength, a well-studied case in quantum metrology [7], strong thermal light, in which case there is little difference between direct and heterodyne detection [8], or even the single-photon state assumed by Gottesman, Jennewein, and Croke in their proposal of shared-entanglement stellar interferometry [9]. This quantum measurement nonlocality can be regarded as a dual of Einstein-Podolsky-Rosen entanglement [4,10]: Despite the fact that bipartite thermal light has a well-defined classical description and possesses no quantum entanglement, nonlocal quantum measurements are necessary to extract the most information from the light. For optical interferometry and imaging applications in general, the result demonstrates the fundamental advantage of nonlocal measurements for weak thermal light and motivates the development of coherent optical measurement techniques, such as integrated optical information processing [2,11,12] and entanglement sharing [9].

Consider the estimation of first-order spatial coherence ($g^{(1)}$) between two distant optical modes. In quantum optics, bipartite thermal light is described by the density operator

$$\rho = \int d^2\alpha\, d^2\beta\, \Phi(\alpha, \beta)|\alpha, \beta\rangle\langle\alpha, \beta|, \qquad (1)$$

where $|\alpha, \beta\rangle$ is a coherent state with amplitudes α and β in the two modes and $\Phi(\alpha, \beta)$ is the Sudarshan-Glauber representation [1], given by

$$\Phi(\alpha, \beta) = \frac{1}{\pi^2 \det\Gamma} \exp\left[-(\,\alpha^* \quad \beta^*\,)\Gamma^{-1}\binom{\alpha}{\beta}\right]. \qquad (2)$$

Γ is the mutual coherence matrix:

$$\Gamma \equiv \begin{pmatrix} \Gamma_{aa} & \Gamma_{ab} \\ \Gamma_{ba} & \Gamma_{bb} \end{pmatrix} = \begin{pmatrix} \langle a^\dagger a \rangle & \langle b^\dagger a \rangle \\ \langle a^\dagger b \rangle & \langle b^\dagger b \rangle \end{pmatrix}, \quad (3)$$

and a and b are annihilation operators of the optical modes. The zero-mean Gaussian statistics are a standard assumption for astronomical sources in theoretical optics [1,3]. The positive Φ function indicates that the two modes are classically correlated only and possess no quantum entanglement [13].

Let $\langle a^\dagger a \rangle = \langle b^\dagger b \rangle = \epsilon/2$ for simplicity. For an incoming light with photon-flux spectral density $S(\nu)$ and a relatively narrow detector bandwidth $\Delta\nu$ around a center frequency ν_0, the filtered photon flux is $S(\nu_0)\Delta\nu$. Over the duration of the effective temporal mode $\Delta t \sim 1/\Delta\nu$, $\epsilon = S(\nu_0)\Delta\nu\Delta t \sim S(\nu_0)$ turns out to be independent of the detector bandwidth and a function of the source and the telescope efficiency only. Considering the case $\epsilon \ll 1$, as is common for interferometry with high optical ν_0, the density operator can be approximated in the photon-number basis as

$$\rho = (1 - \epsilon)|0, 0\rangle\langle 0, 0| + \frac{\epsilon}{2}[|0, 1\rangle\langle 0, 1| + |1, 0\rangle\langle 1, 0|$$

$$+ g^*|0, 1\rangle\langle 1, 0| + g|1, 0\rangle\langle 0, 1|] + O(\epsilon^2), \quad (4)$$

where I have defined $\epsilon g/2 \equiv \Gamma_{ab} = \Gamma_{ba}^*$ and $g \equiv g_1 + ig_2$ as the complex degree of coherence with $|g| \leq 1$ [1]. In the following, I neglect the small $O(\epsilon^2)$ terms, assume that ϵ is known, and g_1 and g_2 are the unknown parameters to be estimated. The assumption of a known ϵ should be reasonable, as other noninterferometric imaging methods can be used to estimate the average photon flux and are usually much less sensitive to noise [2]. Otherwise, ϵ should also be regarded as an unknown parameter to be estimated by the interferometer, a complication outside the scope of this Letter.

Any measurement in quantum mechanics can be modeled by a positive operator-valued measure (POVM) $E(y)$ [4,14], which determines the probability of the observation y:

$$P(y|g) = \text{tr}[E(y)\rho]. \quad (5)$$

For example, in the direct detection scheme [Fig. 1(a)], the two optical modes are brought to interfere at a 50-50 beam splitter and the photons at the two output ports are counted. It can be shown by standard quantum optics calculations [8] that the POVM $E(n, m)$ for photon counts n and m are

$$E(0, 0) = |0, 0\rangle\langle 0, 0|, \quad (6)$$

$$E(1, 0) = \tfrac{1}{2}(|1, 0\rangle + e^{-i\delta}|0, 1\rangle)(\langle 1, 0| + e^{i\delta}\langle 0, 1|), \quad (7)$$

$$E(0, 1) = \tfrac{1}{2}(|1, 0\rangle - e^{-i\delta}|0, 1\rangle)(\langle 1, 0| - e^{i\delta}\langle 0, 1|), \quad (8)$$

where δ is an adjustable phase shift on the b mode. The observation probabilities become
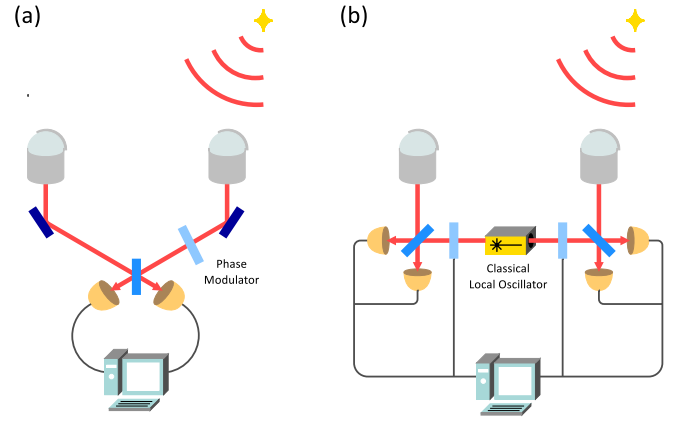

FIG. 1 (color online). Schematics of (a) the direct detection scheme, an example of nonlocal quantum measurement, and (b) a local measurement scheme, which performs spatially separate measurements and permits only classical communication and control between the two sites. Examples of the latter include heterodyne and homodyne detection.

$$P(0, 0|g) = 1 - \epsilon, \quad (9)$$

$$P(1, 0|g) = \frac{\epsilon}{2}[1 + \text{Re}(ge^{-i\delta})], \quad (10)$$

$$P(0, 1|g) = \frac{\epsilon}{2}[1 - \text{Re}(ge^{-i\delta})]. \quad (11)$$

To evaluate the parameter-estimation capability of a measurement scheme, consider the Fisher-information matrix, defined as [15]

$$F \equiv \sum_y \frac{1}{P(y|g)} D(y|g), \quad (12)$$

$$D(y|g) \equiv \begin{pmatrix} \left[\frac{\partial P(y|g)}{\partial g_1}\right]^2 & \frac{\partial P(y|g)}{\partial g_1}\frac{\partial P(y|g)}{\partial g_2} \\ \frac{\partial P(y|g)}{\partial g_2}\frac{\partial P(y|g)}{\partial g_1} & \left[\frac{\partial P(y|g)}{\partial g_2}\right]^2 \end{pmatrix}. \quad (13)$$

The inverse of the Fisher-information matrix provides a lower Cramér-Rao bound to the mean-square estimation error covariance matrix Σ for any unbiased estimate in the form of $\Sigma \geq F^{-1}$. The eigenvalues of F, which must be nonnegative as $F \geq 0$, hence quantify the amounts of independent information obtainable from the measurement. In a total observation time interval T over which the model parameters can be approximated as time constant, $M \sim T/\Delta t \sim T\Delta\nu$ measurements can be performed, and the total Fisher information is $F^{(M)} = MF \sim T\Delta\nu F$. In the limit of large M, the Cramér-Rao bound is asymptotically achievable by maximum-likelihood estimation. This makes the Fisher information a rigorous metric for comparing the inherent capabilities of different measurement schemes for parameter estimation.

The Fisher information for direct detection is

$$F = \frac{\epsilon}{1 - \mathrm{Re}(ge^{-i\delta})^2} \begin{pmatrix} \cos^2\delta & \sin\delta\cos\delta \\ \sin\delta\cos\delta & \sin^2\delta \end{pmatrix}, \quad (14)$$

and the eigenvalues of F are

$$\lambda_1 = 0, \qquad \lambda_2 = \frac{\epsilon}{1 - \mathrm{Re}(ge^{-i\delta})^2}. \quad (15)$$

The zero eigenvalue corresponds to the absence of information about the unobservable quadrature $\mathrm{Im}(ge^{-i\delta})$. In practice, δ is varied over measurements to retrieve information about both quadratures of g. The important point to note here is that $||F|| = \lambda_1 + \lambda_2 \geq \epsilon$ if we take the trace norm. The norm of the total Fisher information for M measurements becomes $||F^{(M)}|| = M||F|| \geq M\epsilon$, which scales linearly with the average photon number $M\epsilon$, thereby achieving the optimal "shot-noise" scaling for parameter estimation using classical states [16]. Similarly, it is shown in [8] that the Fisher information for the shared-entanglement scheme proposed by Gottesman, Jennewein, and Croke [9] has in theory the same expression but reduced by a factor of 2.

Both of the aforementioned schemes can be considered as nonlocal quantum measurements, which require bringing the two modes together physically or sharing entanglement between the two sites. The physical nonlocality makes such schemes increasingly challenging to implement technically as the distance between the two modes increases, primarily due to accumulating decoherence in the form of optical loss along the paths [2]. Local measurement schemes, on the other hand, measure the two modes separately before combining the results via classical communication [Fig. 1(b)], and can therefore be implemented over a much greater distance in principle. To investigate the general performance of any local measurement, let us write the observation probability distribution for POVM $E(y)$ explicitly as

$$P(y|g) = (1 - \epsilon)E_{00,00}(y) + \frac{\epsilon}{2}[E_{01,01}(y) + E_{10,10}(y)$$
$$+ 2|E_{10,01}(y)|\mathrm{Re}(ge^{-i\delta})], \quad (16)$$

where

$$E_{nm,n'm'}(y) \equiv \langle n, m|E(y)|n', m'\rangle \quad (17)$$

and δ is the phase of $E_{10,01}$. To put a bound on the Fisher information given by Eq. (12), note that

$$P(y|g) \geq (1 - \epsilon)E_{00,00}(y), \quad (18)$$

and the positive-semidefinite matrix

$$D = \epsilon^2|E_{10,01}(y)|^2 \begin{pmatrix} \cos^2\delta & \sin\delta\cos\delta \\ \sin\delta\cos\delta & \sin^2\delta \end{pmatrix} \quad (19)$$

defined in Eq. (13) has a trace norm given by $\epsilon^2|E_{10,01}(y)|^2$. Applying the subadditivity property of matrix norms to Eq. (12) results in an upper bound on $||F||$:

$$||F|| \leq \frac{\epsilon^2}{1 - \epsilon}\sum_y \frac{|E_{10,01}(y)|^2}{E_{00,00}(y)}. \quad (20)$$

For generality, I define local measurements as the ones performed using local operations with classical communication (LOCC), which permits the measurement at one site to be conditioned upon the observation at the other site. A necessary condition for a spatial-LOCC POVM is the positive-partial-transpose condition $E^{\mathcal{T}_a}(y) \geq 0$ [17]. By the Cauchy-Schwarz inequality, $|\langle 1, 0|E|0, 1\rangle|^2 = |\langle 0, 0|E^{\mathcal{T}_a}|1, 1\rangle|^2 = |\langle 0, 0|\sqrt{E^{\mathcal{T}_a}}\sqrt{E^{\mathcal{T}_a}}|1, 1\rangle|^2 \leq \langle 0, 0|E^{\mathcal{T}_a} \times |0, 0\rangle\langle 1, 1|E^{\mathcal{T}_a}|1, 1\rangle = \langle 0, 0|E|0, 0\rangle\langle 1, 1|E|1, 1\rangle$, or

$$|E_{10,01}(y)|^2 \leq E_{00,00}(y)E_{11,11}(y). \quad (21)$$

Combining Eqs. (20) and (21), I obtain an $O(\epsilon^2)$ upper bound on $||F||$:

$$||F|| \leq \frac{\epsilon^2}{1 - \epsilon}\sum_y E_{11,11}(y) = \frac{\epsilon^2}{1 - \epsilon}, \quad (22)$$

where $\sum_y E_{11,11}(y) = 1$ comes from the completeness property of a POVM. The neglected $O(\epsilon^2)$ term in the density operator in Eq. (4) contributes an additional $O(\epsilon^2)$ term to P and an $O(\epsilon^3)$ term to D, so the Fisher information would be modified by an $O(\epsilon^3)$ term and the upper bound in Eq. (22) should be rewritten as

$$||F|| \leq \epsilon^2 + O(\epsilon^3). \quad (23)$$

For M measurements, the bound can be generalized to allow for adaptive measurements conditioned upon previous observations, as shown in Ref. [8]:

$$||F^{(M)}|| \leq M[\epsilon^2 + O(\epsilon^3)]. \quad (24)$$

This upper bound shows that the best Fisher information any spatiotemporal-LOCC measurement can achieve is still substantially worse than that of the spatially nonlocal methods ($||F^{(M)}|| \sim M\epsilon$) when $\epsilon \ll 1$. In other words, spatially local measurements are fundamentally much less efficient than nonlocal methods in extracting coherence information from weak-thermal-light interferometry. This general proof is supported by explicit Fisher-information calculations for heterodyne and homodyne detection [8], signal-to-noise-ratio calculations for direct and heterodyne detection of the full thermal state given by Eq. (1) [8], and the known fact in astronomy that direct detection performs better than heterodyne detection for high optical ν_0 [3,6]. Reference [8] also includes a discussion of the quantum origin of the nonlocality in terms of the semiclassical photodetection picture.

Note that the advantage of nonlocal measurements is lost for coherent states, strong thermal light with $\epsilon \gg 1$ [8], or

even the nonclassical single-photon state studied in Ref. [9]. For coherent states, $|g| = 1$ and the unknown parameters are the phases of the two optical modes in a product of coherent states, in which case it can easily be shown that nonlocal measurements are not necessary, analogous to the case of single-parameter phase estimation with a product state [18]. For strong thermal light with $\epsilon \gg 1$, calculations in Ref. [8] show that the performances of direct detection and heterodyne detection converge and suggest that the noise in this regime is dominated by the thermal statistics of the source rather than the detection statistics. The single-photon state studied in Ref. [9] can also be analyzed using the formalism here by omitting $O(\epsilon^2)$ terms and then putting $\epsilon = 1$, resulting in comparable performances for local and nonlocal measurements.

The peculiar existence of quantum nonlocality for weak thermal light, as a property of bipartite measurements applied to certain separable states, can be regarded as a dual of Einstein-Podolsky-Rosen entanglement [4,10], a property of bipartite states that can produce higher correlations in certain separable measurements. In the context of quantum communication theory, it is well known that nonlocal measurements can extract more information from states with no entanglement [4,5,19]; the result here provides a striking example in which the same type of quantum nonlocality readily exists for observers extracting information from nature.

For practical applications, the result here demonstrates the fundamental advantage of nonlocal quantum measurements for weak-thermal-light interferometry and may have further implications for optical imaging systems, such as compound-eye imaging and fluorescence microscopy [11]. The shared-entanglement proposal in Ref. [9] requires a path-entangled single-photon source and quantum repeaters, both of which are unlikely to become feasible in the near future, but standard linear optics can also perform nonlocal measurements by coherently processing multiple optical modes before detection, provided that optical loss can be minimized. In the short term, the result here thus motivates the development of low-loss coherent optical information devices, such as photonic crystal fibers and integrated photonics, for thermal-light interferometry and imaging [2,11,12].

Accurate coherence information can be obtained only in the limit of many collected photons. This corresponds to measurements of many copies of the quantum state. A more general quantum measurement strategy than the ones considered here involves joint quantum operations on the multiple copies before measurements. This kind of temporal nonlocality is not needed for parameter estimation when spatially nonlocal measurements can be performed [16]. It remains an interesting open question whether coherent temporally nonlocal strategies can offer any significant advantage when one is restricted to spatially local measurements. Other potential generalizations include time-varying parameters and the estimation of temporal coherence for spectroscopy in addition to spatial coherence. One must then take into account the dynamics of the source and colored noise, which can be analyzed using the quantum waveform estimation framework developed in Refs. [20].

*eletmk@nus.edu.sg

[1] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995).

[2] J. D. Monnier, Rep. Prog. Phys. **66**, 789 (2003).

[3] A. R. Thompson, J. M. Moran, and G. W. Swenson Jr., *Interferometry and Synthesis in Radio Astronomy* (Wiley-VCH, Weinheim, 2001).

[4] A. S. Holevo, *Statistical Structure of Quantum Theory* (Springer-Verlag, Berlin, 2001).

[5] A. Peres and W. K. Wootters, Phys. Rev. Lett. **66**, 1119 (1991); C. H. Bennett *et al.*, Phys. Rev. A **59**, 1070 (1999).

[6] C. H. Townes, in *Principles of Long Baseline Stellar Interferometry*, edited by Peter R. Lawson, NASA Publication No. 00-009 (NASA-JPL, Pasadena, CA, 2000), Chap. 4, pp. 59–70, http://olbin.jpl.nasa.gov/iss1999/coursenotes.html.

[7] C. M. Caves and A. Shaji, Opt. Commun. **283**, 695 (2010).

[8] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.107.270402 for supporting calculations and discussions.

[9] D. Gottesman, T. Jennewein, and S. Croke, arXiv:1107.2939.

[10] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[11] D. J. Brady, *Optical Imaging and Spectroscopy* (Wiley, Hoboken, 2009).

[12] D. Psaltis, Science **298**, 1359 (2002).

[13] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).

[14] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976); H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2010).

[15] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I* (Wiley, New York, 2001).

[16] V. Giovannetti, S. Lloyd, and L. Maccone, Nature Photon. **5**, 222 (2011).

[17] B. M. Terhal, D. P. DiVincenzo, and D. W. Leung, Phys. Rev. Lett. **86,** 5807 (2001).

[18] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96,** 010401 (2006).

[19] S. Lloyd, Science **321,** 1463 (2008); S. H. Tan *et al.*, Phys. Rev. Lett. **101,** 253601 (2008); S. Guha and B. I. Erkmen, Phys. Rev. A **80,** 052310 (2009); S. Pirandola, Phys. Rev. Lett. **106,** 090504 (2011).

[20] M. Tsang, Phys. Rev. Lett. **102,** 250403 (2009); M. Tsang, H. M. Wiseman, and C. M. Caves, Phys. Rev. Lett. **106,** 090401 (2011).

Quantum Nonlocality in Weak-Thermal-Light Interferometry: Supplementary Material

Mankei Tsang*

Department of Electrical and Computer Engineering,

National University of Singapore, 4 Engineering Drive 3, Singapore 117583

Department of Physics, National University of Singapore,

2 Science Drive 3, Singapore 117551 and

Center for Quantum Information and Control,

University of New Mexico, MSC07–4220,

Albuquerque, New Mexico 87131-0001, USA

This Supplementary Material contains supportive calculations and discussions that complement the main text. Section I derives the positive operator-valued measure (POVM) for direct detection, Sec. II derives the POVM for shared-entanglement interferometry and the resulting Fisher information, Sec. III calculates a bound on the total Fisher information for multiple adaptive measurements, Secs. IV and V calculate the Fisher information for heterodyne and homodyne detection, Sec. VI investigates the performances of direct and heterodyne detection for strong thermal light, and Sec. VII discusses the origin of the quantum nonlocality in terms of the semiclassical photodetection picture. The reference list at the end is identical to the one in the main text for easier cross-referencing.

I. POVM FOR DIRECT DETECTION

In direct detection, the optical modes a and b are combined by a beam splitter and photon-counting is performed at the two output ports. Let U be the unitary operator that corresponds to the operation of the beam splitter on the bipartite quantum state ρ. The observation probability distribution is

$$P(n,m|g) = \langle n,m|U\rho U^\dagger|n,m\rangle, \tag{S1}$$

---
*Electronic address: eletmk@nus.edu.sg

where $|n, m\rangle$ is a Fock state. We can then write the POVM as

$$E(n, m) = U^\dagger |n, m\rangle\langle n, m| U, \tag{S2}$$

which propagates the Fock-state projection back to the time when the state of light is ρ.

With at most one photon in the quantum state, we are interested in $(n, m) = (0, 0), (1, 0), (0, 1)$ only. Applying the unitary to the Fock states,

$$U^\dagger |0, 0\rangle = |0, 0\rangle, \tag{S3}$$

$$U^\dagger |1, 0\rangle = U^\dagger a^\dagger U U^\dagger |0, 0\rangle = U^\dagger a^\dagger U |0, 0\rangle \tag{S4}$$

$$= \frac{1}{\sqrt{2}} \left(a^\dagger + e^{-i\delta} b^\dagger\right) |0, 0\rangle \tag{S5}$$

$$= \frac{1}{\sqrt{2}} \left(|1, 0\rangle + e^{-i\delta}|0, 1\rangle\right), \tag{S6}$$

$$U^\dagger |0, 1\rangle = \frac{1}{\sqrt{2}} \left(|1, 0\rangle - e^{-i\delta}|0, 1\rangle\right). \tag{S7}$$

The POVM is hence

$$E(0, 0) = |0, 0\rangle\langle 0, 0|, \tag{S8}$$

$$E(1, 0) = \frac{1}{2} \left(|1, 0\rangle + e^{-i\delta}|0, 1\rangle\right) \left(\langle 1, 0| + e^{i\delta}\langle 0, 1|\right), \tag{S9}$$

$$E(0, 1) = \frac{1}{2} \left(|1, 0\rangle - e^{-i\delta}|0, 1\rangle\right) \left(\langle 1, 0| - e^{i\delta}\langle 0, 1|\right). \tag{S10}$$

II. SHARED-ENTANGLEMENT INTERFEROMETRY

Assuming an entangled ancilla in two modes c and d given by

$$|\delta\rangle \equiv \frac{1}{\sqrt{2}} \left(|0, 1\rangle_{c,d} + e^{i\delta}|1, 0\rangle_{c,d}\right), \tag{S11}$$

with each mode sent to the sites of a and b modes for separate interference measurements [9], the POVM for photon counts (n, m, n', m') is

$$E(n, m, n', m') = \langle\delta|U_{ac}^\dagger \otimes U_{bd}^\dagger|n, m, n', m'\rangle\langle n, m, n', m'|U_{ac} \otimes U_{bd}|\delta\rangle, \tag{S12}$$

where U_{ac} denotes the beam-splitting unitary on modes a and c and U_{bd} denotes the same unitary on modes b and d. The calculation is more involved but similar to the one for direct

detection. The final result is

$$E(y_0) = |0,0\rangle\langle 0,0|, \tag{S13}$$

$$E(y_1) = \frac{1}{2}|0,1\rangle\langle 0,1|, \tag{S14}$$

$$E(y_2) = \frac{1}{2}|1,0\rangle\langle 1,0|, \tag{S15}$$

$$E(y_3) = \frac{1}{4}\left(|1,0\rangle + e^{-i\delta}|0,1\rangle\right)\left(\langle 1,0| + e^{i\delta}\langle 0,1|\right), \tag{S16}$$

$$E(y_4) = \frac{1}{4}\left(|1,0\rangle - e^{-i\delta}|0,1\rangle\right)\left(\langle 1,0| - e^{i\delta}\langle 0,1|\right), \tag{S17}$$

where each y_j corresponds to a set of (n, m, n', m') that produce the same POVM. When applied to the quantum state ρ, only observations y_3 and y_4 contribute to the Fisher information about g. Since $E(y_3) = E(1,0)/2$ and $E(y_4) = E(0,1)/2$, the Fisher information for shared-entanglement interferometry is simply that for direct detection reduced by a factor of 2:

$$F = \frac{\epsilon}{2[1 - \text{Re}(ge^{-i\delta})^2]}\begin{pmatrix} \cos^2\delta & \sin\delta\cos\delta \\ \sin\delta\cos\delta & \sin^2\delta \end{pmatrix}. \tag{S18}$$

III. ADAPTIVE MEASUREMENTS

For M measurements, the joint observation probability distribution can be written as

$$P(y_M, \ldots, y_1|g) = P(y_M|g, y_{M-1}, \ldots, y_1)P(y_{M-1}, \ldots, y_1|g). \tag{S19}$$

Each element of the total Fisher information matrix for M measurements, using an alternate form of the Fisher matrix [15], becomes

$$F_{jk}^{(M)} \equiv -\sum_{y_1,\ldots,y_M} P(y_M, \ldots, y_1|g)\frac{\partial^2}{\partial g_j \partial g_k}\ln P(y_M, \ldots, y_1|g) \tag{S20}$$

$$= -\sum_{y_1,\ldots,y_M} P(y_M|g, y_{M-1}, \ldots, y_1)P(y_{M-1}, \ldots, y_1|g)$$

$$\times \left[\frac{\partial^2}{\partial g_j \partial g_k}\ln P(y_M|g, y_{M-1}, \ldots, y_1) + \frac{\partial^2}{\partial g_j \partial g_k}\ln P(y_{M-1}, \ldots, y_1|g)\right] \tag{S21}$$

$$= \sum_{y_1,\ldots,y_{M-1}} P(y_{M-1}, \ldots, y_1|g)F_{Mjk}(y_{M-1}, \ldots, y_1) + F_{jk}^{(M-1)}, \tag{S22}$$

where F_M denotes the conditional Fisher information of the Mth measurement:

$$F_{Mjk}(y_{M-1}, \ldots, y_1) \equiv -\sum_{y_M} P(y_M|g, y_{M-1}, \ldots, y_1)\frac{\partial^2}{\partial g_j \partial g_k}\ln P(y_M|g, y_{M-1}, \ldots, y_1). \tag{S23}$$

Applying the subadditivity property of matrix norms,

$$||F^{(M)}|| \leq \sum_{y_1,\ldots,y_{M-1}} P(y_{M-1},\ldots,y_1|g)||F_M|| + ||F^{(M-1)}|| \tag{S24}$$

$$\leq \max_{y_1,\ldots,y_{M-1}} ||F_M|| + ||F^{(M-1)}||, \tag{S25}$$

and by induction,

$$||F^{(M)}|| \leq \sum_{m=1}^{M} \max_{y_1,\ldots,y_{m-1}} ||F_m||. \tag{S26}$$

This proves that the norm of the total Fisher information cannot exceed the sum of the maximized single-measurement values.

For the mth quantum measurement with outcome y_m conditioned upon previous observations, we can write

$$P(y_m|g, y_{m-1},\ldots,y_1) = \text{tr}\left[E(y_m|y_{m-1},\ldots,y_1)\rho\right]. \tag{S27}$$

This means that the bound given by Eq. (23) in the main text for spatial-LOCC measurements in the case of $\epsilon \ll 1$ is also applicable to $||F_m||$:

$$||F_m|| \leq \epsilon^2 + O(\epsilon^3). \tag{S28}$$

The total Fisher information is hence bounded by

$$||F^{(M)}|| \leq M\left[\epsilon^2 + O(\epsilon^3)\right], \tag{S29}$$

which generalizes the bound to the case of multiple spatiotemporal-LOCC measurements and proves that no adaptive strategy can improve the scaling $||F^{(M)}|| \sim M\epsilon^2$.

IV. HETERODYNE DETECTION

The POVM for heterodyne detection is [14]

$$E(\mu, \nu) = \frac{1}{\pi^2}|\mu, \nu\rangle\langle\mu, \nu|, \tag{S30}$$

where $|\mu, \nu\rangle$ is a coherent state and the normalization is $\int d^2\mu d^2\nu E(\mu, \nu) = I$, the identity operator. The relevant POVM matrix elements are

$$E_{00,00}(\mu, \nu) \equiv \frac{1}{\pi^2} |\langle 0, 0|\mu, \nu\rangle|^2 \tag{S31}$$

$$= \frac{1}{\pi^2} \exp\left(-|\mu|^2 - |\nu|^2\right), \tag{S32}$$

$$E_{01,01}(\mu, \nu) \equiv \frac{1}{\pi^2} |\langle 0, 1|\mu, \nu\rangle|^2 \tag{S33}$$

$$= \frac{1}{\pi^2} \exp\left(-|\mu|^2 - |\nu|^2\right) |\nu|^2, \tag{S34}$$

$$E_{10,10}(\mu, \nu) \equiv \frac{1}{\pi^2} |\langle 1, 0|\mu, \nu\rangle|^2 \tag{S35}$$

$$= \frac{1}{\pi^2} \exp\left(-|\mu|^2 - |\nu|^2\right) |\mu|^2, \tag{S36}$$

$$E_{01,10}(\mu, \nu) \equiv \frac{1}{\pi^2} \langle 0, 1|\mu, \nu\rangle\langle\mu, \nu|1, 0\rangle \tag{S37}$$

$$= \frac{1}{\pi^2} \exp\left(-|\mu|^2 - |\nu|^2\right) \mu^*\nu. \tag{S38}$$

The Fisher information is hence

$$F = \frac{\epsilon^2}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + O(\epsilon^3), \tag{S39}$$

$$||F|| = \epsilon^2 + O(\epsilon^3). \tag{S40}$$

This shows that the performance of heterodyne detection is already the optimum allowed by quantum mechanics for any local measurement according to the bound given by Eq. (23) in the main text.

V. HOMODYNE DETECTION

For homodyne detection,

$$E(x, y) = |x, y\rangle\langle x, y|, \tag{S41}$$

where $|x, y\rangle$ is a quadrature eigenstate:

$$\frac{1}{\sqrt{2}} \left(ae^{-i\delta_a} + a^\dagger e^{i\delta_a}\right) |x, y\rangle = x|x, y\rangle, \tag{S42}$$

$$\frac{1}{\sqrt{2}} \left(be^{-i\delta_b} + b^\dagger e^{i\delta_b}\right) |x, y\rangle = y|x, y\rangle, \tag{S43}$$

δ_a and δ_b are local-oscillator phases, and the normalization is $\int dx dy E(x, y) = I$. The relevant POVM elements are

$$E_{00,00}(x, y) = \frac{1}{\pi} \exp\left(-x^2 - y^2\right), \tag{S44}$$

$$E_{01,01}(x, y) = \frac{2}{\pi} \exp\left(-x^2 - y^2\right) y^2, \tag{S45}$$

$$E_{10,10}(x, y) = \frac{2}{\pi} \exp\left(-x^2 - y^2\right) x^2, \tag{S46}$$

$$E_{10,01}(x, y) = \frac{2}{\pi} e^{i\delta} \exp\left(-x^2 - y^2\right) xy, \tag{S47}$$

where $\delta \equiv \delta_a - \delta_b$. The Fisher information becomes

$$F = \epsilon^2 \begin{pmatrix} \cos^2 \delta & \sin \delta \cos \delta \\ \sin \delta \cos \delta & \sin^2 \delta \end{pmatrix} + O(\epsilon^3), \tag{S48}$$

$$||F|| = \epsilon^2 + O(\epsilon^3). \tag{S49}$$

Homodyne detection is also able to saturate the bound given by Eq. (23) in the main text, but each measurement gives information about only one quadrature of g and δ should be varied over measurements to estimate both quadratures.

Qualitatively, the inferior Fisher information for heterodyne and homodyne detection can be attributed to the non-zero vacuum fluctuations even when no photon is coming in to provide information about the unknown parameters. Nonlocal measurements are able to perfectly discriminate against this case and discard the useless observations, but heterodyne or homodyne detection is unable to do so and forced to include vacuum fluctuations as potentially useful observations, resulting in a substantially worse estimation accuracy in the long run.

VI. THERMAL LIGHT WITH ARBITRARY ϵ

For $\epsilon \gtrsim 1$, it is necessary to use the full thermal state given by Eq. (1) in the main text. First consider the observation probability density of heterodyne detection:

$$P(\mu, \nu|g) = \int d^2\alpha d^2\beta \Pi(\mu, \nu|\alpha, \beta) \Phi(\alpha, \beta|g), \tag{S50}$$

$$\Pi(\mu, \nu|\alpha, \beta) \equiv \langle \alpha, \beta | E(\mu, \nu) | \alpha, \beta \rangle \tag{S51}$$

$$= \frac{1}{\pi^2} \exp\left(-|\mu - \alpha|^2 - |\nu - \beta|^2\right). \tag{S52}$$

We can interpret these expressions using a semiclassical photodetection picture [1]: The heterodyne detection statistics obey $\Pi(\mu, \nu | \alpha, \beta)$ for given classical fields (α, β), but the fields from the source also have a statistical distribution given by $\Phi(\alpha, \beta)$, so the marginal observation density is taken to be Π averaged over Φ. The resulting convolution of the two Gaussians can be calculated analytically and given by

$$P(\mu, \nu | g) = \frac{1}{\pi^2 \det \Gamma'} \exp \left[- \begin{pmatrix} \mu^* & \nu^* \end{pmatrix} \Gamma'^{-1} \begin{pmatrix} \mu \\ \nu \end{pmatrix} \right], \tag{S53}$$

where the new covariance matrix Γ' is

$$\Gamma' = \Gamma + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \epsilon/2 + 1 & \epsilon g/2 \\ \epsilon g^*/2 & \epsilon/2 + 1 \end{pmatrix}. \tag{S54}$$

The factors of 1 come from Π and represent detection noise. When $\epsilon \gg 1$, Π is much sharper than Φ, so the convolution essentially reproduces Φ as the marginal observation density and $P(\mu, \nu | g) \approx \Phi(\mu, \nu | g)$. In other words, the inherent thermal noise from the source overwhelms the heterodyne detection noise in the $\epsilon \gg 1$ regime.

To estimate the performance of heterodyne detection, we can consider the signal-to-noise ratio (SNR) [3, 6]. If we take $\mu\nu^*$ as the output signal, $\langle \mu\nu^* \rangle = \epsilon g/2$ directly gives g on average, and the signal energy is

$$S \equiv |\langle \mu\nu^* \rangle|^2 = \frac{\epsilon^2 |g|^2}{4}. \tag{S55}$$

The noise energy is

$$N \equiv \langle |\mu\nu^*|^2 \rangle - S. \tag{S56}$$

The fourth-order field statistics can be computed with the help of the matrix $G \equiv \Gamma'^{-1}$:

$$\det G = G_{aa} G_{bb} - G_{ab} G_{ba}, \tag{S57}$$

$$\langle |\mu\nu^*|^2 \rangle = \det G \frac{\partial^2}{\partial G_{aa} \partial G_{bb}} \frac{1}{\det G} \tag{S58}$$

$$= \frac{2 G_{aa} G_{bb}}{\det G^2} - \frac{1}{\det G} \tag{S59}$$

$$= \left(1 + \frac{\epsilon}{2} \right)^2 + \frac{\epsilon^2 |g|^2}{4}, \tag{S60}$$

$$N = \left(1 + \frac{\epsilon}{2} \right)^2. \tag{S61}$$

The SNR is hence

$$\frac{S}{N} = \frac{\epsilon^2 |g|^2}{(2 + \epsilon)^2}. \tag{S62}$$

For $\epsilon \ll 1$, $S/N \approx \epsilon^2 |g|^2/4$, but for $\epsilon \gg 1$, the SNR saturates to $S/N \to |g|^2$ and becomes independent of ϵ.

For direct detection,

$$P(n, m|g) = \int d^2\alpha \, d^2\beta \, \Pi(n, m|\alpha, \beta) \Phi(\alpha, \beta|g), \tag{S63}$$

$$\Pi(n, m|\alpha, \beta) \equiv \langle \alpha, \beta | E(n, m) | \alpha, \beta \rangle = |\langle n, m | u, v \rangle|^2, \tag{S64}$$

$$\begin{pmatrix} u \\ v \end{pmatrix} \equiv V \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad V \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & e^{i\delta} \\ 1 & -e^{i\delta} \end{pmatrix}. \tag{S65}$$

Changing the integration variables to (u, v),

$$P(n, m|g) = \int d^2u \, d^2v \, \Pi'(n, m|u, v) \Phi'(u, v|g), \tag{S66}$$

$$\Pi'(n, m|u, v) \equiv \exp(-|u|^2) \frac{|u|^{2n}}{n!} \exp(-|v|^2) \frac{|v|^{2m}}{m!}, \tag{S67}$$

$$\Phi'(u, v|g) \equiv \frac{\det K}{\pi^2} \exp \left[-\begin{pmatrix} u^* & v^* \end{pmatrix} K \begin{pmatrix} u \\ v \end{pmatrix} \right], \tag{S68}$$

$$K \equiv V \Gamma^{-1} V^\dagger = \frac{2}{\epsilon(1 - |g|^2)} \begin{pmatrix} 1 - \mathrm{Re}(g e^{-i\delta}) & i \, \mathrm{Im}(g e^{-i\delta}) \\ -i \, \mathrm{Im}(g e^{-i\delta}) & 1 + \mathrm{Re}(g e^{-i\delta}) \end{pmatrix}, \tag{S69}$$

$$\det K = K_{aa} K_{bb} - K_{ab} K_{ba} = \frac{4}{\epsilon^2(1 - |g|^2)}. \tag{S70}$$

The averaging of a Poissonian with a Gaussian is difficult to calculate exactly. For $\epsilon \gg 1$, however, the photon counts (n, m) should be large most of the time, and $P(n, m|g)$ may be approximated by a Gaussian. Let us therefore focus on the first and second moments of (n, m) for $P(n, m|g)$. The first moment of n is

$$\langle n \rangle = \sum_{n,m} n P(n, m|g) \tag{S71}$$

$$= \int d^2u \, d^2v \, |u|^2 \Phi'(u, v|g) \tag{S72}$$

$$= -\det K \frac{\partial}{\partial K_{aa}} \frac{1}{\det K}. \tag{S73}$$

Similarly,

$$\langle m \rangle = -\det K \frac{\partial}{\partial K_{bb}} \frac{1}{\det K}, \tag{S74}$$

$$\langle n^2 \rangle = \langle n \rangle + \det K \frac{\partial^2}{\partial K_{aa}^2} \frac{1}{\det K}, \tag{S75}$$

$$\langle m^2 \rangle = \langle m \rangle + \det K \frac{\partial^2}{\partial K_{bb}^2} \frac{1}{\det K}, \tag{S76}$$

$$\langle nm \rangle = \det K \frac{\partial^2}{\partial K_{aa} \partial K_{bb}} \frac{1}{\det K}. \tag{S77}$$

This gives

$$\langle n \rangle = \frac{\epsilon}{2} \left[1 + \mathrm{Re}(ge^{-i\delta}) \right], \tag{S78}$$

$$\langle m \rangle = \frac{\epsilon}{2} \left[1 - \mathrm{Re}(ge^{-i\delta}) \right], \tag{S79}$$

$$\langle \Delta n^2 \rangle = \langle n \rangle + \langle n \rangle^2, \tag{S80}$$

$$\langle \Delta m^2 \rangle = \langle m \rangle + \langle m \rangle^2, \tag{S81}$$

$$\langle \Delta n \Delta m \rangle = \frac{\epsilon^2}{4} \mathrm{Im}(ge^{-i\delta})^2. \tag{S82}$$

A behavior similar to the case of heterodyne detection can be seen here. For $\langle n \rangle, \langle m \rangle \sim \epsilon \gg 1$, the noise covariances scale as ϵ^2 rather than ϵ, indicating that the source thermal noise also overwhelms the Poissonian detection noise.

Since the observation statistics are expected to be approximately Gaussian for $\epsilon \gg 1$, we can similarly consider the SNR as a performance metric. Taking the output signal as $n - m$, the average of which gives $\langle n - m \rangle = \epsilon \, \mathrm{Re}(ge^{-i\delta})$, a quadrature of g, the signal energy is

$$S = \langle n - m \rangle^2 = \epsilon^2 \, \mathrm{Re}(ge^{-i\delta})^2, \tag{S83}$$

and the noise energy is

$$N = \langle (n - m)^2 \rangle - S \tag{S84}$$

$$= \langle \Delta n^2 \rangle + \langle \Delta m^2 \rangle - 2 \langle \Delta n \Delta m \rangle \tag{S85}$$

$$= \epsilon + \frac{\epsilon^2}{2} \left[1 + \mathrm{Re}(ge^{-i\delta})^2 - \mathrm{Im}(ge^{-i\delta})^2 \right]. \tag{S86}$$

If we perform two measurements, one with $\delta = \delta_1$ and one with $\delta = \delta_1 + \pi/2$ to measure the other quadrature of g, the average signal and noise energies per measurement becomes

$$\bar{S} = \frac{\epsilon^2 |g|^2}{2}, \tag{S87}$$

$$\bar{N} = \epsilon + \frac{\epsilon^2}{2}, \tag{S88}$$

and the average SNR is

$$\frac{\bar{S}}{\bar{N}} = \frac{\epsilon|g|^2}{2+\epsilon}. \tag{S89}$$

For $\epsilon \gg 1$, the SNR saturates to $|g|^2$, just like the SNR of heterodyne detection, suggesting that the SNR is dominated by source thermal noise regardless of the detection method and nonlocal measurements do not have an advantage when $\epsilon \gg 1$.

For $\epsilon \ll 1$, the SNR is still a valid performance metric for a large number of measurements, in which case the statistics become approximately Gaussian by the central limit theorem and averaging M observations improves the final SNR by a factor of M. The direct-detection SNR is $\approx M\epsilon|g|^2/2$ and significantly better than the heterodyne SNR $\approx M\epsilon^2|g|^2/4$, a fact well known in astronomy [3, 6] and rigorously generalized in this paper.

VII.   QUANTUM ORIGIN OF MEASUREMENT NONLOCALITY IN THE SEMICLASSICAL PHOTODETECTION PICTURE

One may well wonder where quantum mechanics comes in, if both Φ and Π are nonnegative and the whole problem obeys classical statistics in the semiclassical photodetection picture. The answer lies in the fact that the likelihood function

$$\Pi(y|\alpha,\beta) \equiv \langle\alpha,\beta|E(y)|\alpha,\beta\rangle \tag{S90}$$

cannot be an arbitrarily sharp probability distribution in quantum mechanics. It is the Husimi representation [1], more commonly applied to a quantum state but here to a POVM.

If we regard $\Pi(y|\alpha,\beta)$ as a likelihood function of (α,β) for a given observation y, the sharpness of $\Pi(y|\alpha,\beta)$ with respect to (α,β) in phase space characterizes the amount of information about (α,β) contained in the observation y. The Husimi representation has a maximum magnitude and a finite variance for each quadrature, which means that there is a limited amount of information about the fields that an observation can provide.

The information of mutual coherence lies only in the nonlocal second-order field correlation $\alpha\beta^*$ for thermal light, the first-order mean fields of which are zero. If $E(y)$ corresponds to a local measurement and is separable into $E_a(y) \otimes E_b(y)$, the sharpness of $\Pi(y|\alpha,\beta) = \Pi_a(y|\alpha)\Pi_b(y|\beta)$ with respect to $\alpha\beta^*$ would be more limited than that allowed by nonlocal measurements, meaning that local measurements extract less information about

the coherence than nonlocal measurements. In this sense, the measurement nonlocality can be regarded as a dual of Einstein-Podolsky-Rosen entanglement [4, 10]; the former is a property of bipartite measurements that can extract more information from certain separable states and the latter a property of bipartite states that can produce higher correlations in certain separable measurements.

---

[1] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995).

[2] J. D. Monnier, Rep. Prog. Phys. **66**, 789 (2003).

[3] A. R. Thompson, J. M. Moran, and G. W. Swenson Jr. *Interferometry and Synthesis in Radio Astronomy* (Wiley-VCH, Weinheim, 2001).

[4] A. S. Holevo, *Statistical Structure of Quantum Theory* (Springer-Verlag, Berlin, 2001).

[5] A. Peres and W. K. Wootters, Phys. Rev. Lett. **66**, 1119 (1991); C. H. Bennett *et al.*, Phys. Rev. A **59**, 1070 (1999).

[6] C. H. Townes, Noise and sensitivity in interferometry. Ch. 4 in *Principles of Long Baseline Stellar Interferometry*, ed. Peter R. Lawson, NASA-JPL (Publication 00-009), pp. 59-70 (2000), `http://olbin.jpl.nasa.gov/iss1999/coursenotes.html`.

[7] C. M. Caves and A. Shaji, Opt. Commun. **283**, 695 (2010).

[8] See Supplementary Material for supporting calculations and discussions.

[9] D. Gottesman, T. Jennewein, and S. Croke, e-print arXiv:1107.2939.

[10] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[11] D. J. Brady, *Optical Imaging and Spectroscopy* (Wiley, Hoboken, 2009).

[12] D. Psaltis, Science **298**, 1359 (2002).

[13] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).

[14] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976); H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2010).

[15] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I* (Wiley, New York, 2001).

[16] V. Giovannetti, S. Lloyd, and L. Maccone, Nature Photon. **5**, 222 (2011).

[17] B. M. Terhal, D. P. DiVincenzo, and D. W. Leung, Phys. Rev. Lett. **86**, 5807 (2001).

[18] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).

[19] S. Lloyd, Science **321**, 1463 (2008); S. H. Tan *et al.*, Phys. Rev. Lett. **101**, 253601 (2008); S. Guha and B. I. Erkmen, Phys. Rev. A **80**, 052310 (2009); S. Pirandola, Phys. Rev. Lett. **106**, 090504 (2011).

[20] M. Tsang, Phys. Rev. Lett. **102**, 250403 (2009); M. Tsang, H. M. Wiseman, and C. M. Caves, Phys. Rev. Lett. **106**, 090401 (2011).